

Wright State University

CORE Scholar

International Symposium on Aviation
Psychology - 2007

International Symposium on Aviation
Psychology

2007

Effects of Imperfect Automation Support on Air Traffic Controller (ATCO) Performance, Mental Workload, and Attention Allocation: Miss vs. False Alarm Prone Automation

Ericka Rovira

Raja Parasuraman

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2007



Part of the [Other Psychiatry and Psychology Commons](#)

Repository Citation

Rovira, E., & Parasuraman, R. (2007). Effects of Imperfect Automation Support on Air Traffic Controller (ATCO) Performance, Mental Workload, and Attention Allocation: Miss vs. False Alarm Prone Automation. *2007 International Symposium on Aviation Psychology*, 611-617.
https://corescholar.libraries.wright.edu/isap_2007/31

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2007 by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

EFFECTS OF IMPERFECT AUTOMATION SUPPORT ON AIR TRAFFIC CONTROLLER (ATCo) PERFORMANCE, MENTAL WORKLOAD, AND ATTENTION ALLOCATION: MISS VS. FALSE ALARM PRONE AUTOMATION

Ericka Rovira and Raja Parasuraman
George Mason University
Fairfax, VA

This study investigated ATCo performance with reliable and imperfect automation support to examine whether automation benefits would accrue in a mixed equipage environment where ATCos were no longer responsible for freely maneuvering aircraft. Of particular interest was how types of automation imperfections would affect ATCo performance and attention allocation. Further, it was studied if automation imperfection would affect ATCo trust and attention allocation to freely maneuvering aircraft. With reliable automation ATCos detected conflicts better and sooner, compared to manual performance. Second, with imperfect automation primary task performance degraded significantly and ATCos felt more self-confident to perform without the automation when they were supported with imperfect compared to reliable automation. With respect to types of imperfection, there was a marginal trend for worse performance with miss prone automation as compared to false alarm prone automation. Third, while there were no differences in trust ratings or eye movements to freely maneuvering aircraft, ATCos perceived that freely maneuvering aircraft improved their performance more with automation than without.

Introduction

Previous research investigating automation imperfection in a shared decision making system, suggests that imperfect automation may not be able to adequately support ATCos when they are left to monitor a system in which pilots are making the separation decisions (Metzger & Parasuraman, 2005). If ATCos are still required to step-in and resolve a conflict in instances where pilots fail to successfully obtain a resolution in a given amount of time, a heavy monitoring load may be imposed on ATCos, particularly if traffic density were increased and the time to detect and resolve a conflict were reduced (Metzger & Parasuraman, 2001). This is problematic because performance could potentially degrade not only because automation moves the ATCo further away from the decision-making loop, but because ATCos may become passive monitors of pilot actions.

A contrary view is that controllers need not be responsible for aircraft that have the onboard technology to safely separate themselves from other aircraft. If pilots had full responsibility for their own separation from other aircraft, then ATCos would be freed of monitoring freely maneuvering aircraft and could provide better service to managed aircraft when there are many aircraft in the airspace. To date, there is little supporting evidence that ATCos can appropriately disengage their attention from the freely maneuvering aircraft when there are many aircraft in the airspace. Initially, only some aircraft will have the equipment to freely maneuver, whereas other less well-equipped aircraft will not. This is the *mixed-equipage* issue. This study will provide the

first empirical evidence in a mixed-equipage environment where ATCos are not responsible for detecting self-separations of freely maneuvering aircraft. Perhaps because ATCos were responsible for detecting self-separations for freely maneuvering aircraft in previous studies, workload increased (Corker, et al., 2003) and the assumed workload reductions were not found (Metzger, et al., 2003).

Additionally, there is a growing body of research investigating the differential impact of types of automation imperfection (miss versus false alarm) on human automation interaction (Meyer, 2001; 2004; Parasuraman & Riley, 1997). The assumption is that depending on the type of automation imperfection (miss or false alarm) operators' reliance and compliance to an automated system may be affected by the imperfection. Previous research has shown that miss prone automation was correlated with poorer performance on concurrent tasks and false alarm prone automation was correlated with poorer performance on the automated task (Dixon & Wickens, 2006). Dixon and Wickens (2006) findings support the reliance-compliance distinction such that miss prone automation resulted in increased visual attention to the system monitoring task at a cost to the other tasks, and false alarm prone automation resulted in slowed alert driven shifts of visual attention to the system monitoring task. To date, no studies have investigated the effects of these types of automation imperfections on operator performance when the *primary task* is automated, but not 100% reliable.

Research Questions and Hypotheses

The purpose of this experiment was to investigate the effects of reliable and imperfect automation on ATCo performance, mental workload, and attention allocation. We also examined whether ATCos would show complacency effects when using automation that was not perfectly reliable. It was expected that ATCos would detect more conflicts, detect conflicts sooner, and show better performance in concurrent task performance when they were supported by automation as opposed to when they were performing the tasks manually. Therefore, the detection of conflicts with automation support was compared to the detection of the same conflicts when the ATCos were manually detecting them. Mental workload was thought to be reduced when ATCos were supported by automation. Additionally, if complacency is indeed due to high trust and reduced attention allocation (Parasuraman et al., 1993) and if attention and eye movements are related to some degree, then it was expected that with automation ATCos would reduce the number of fixations to the radar display compared to manual control.

By automation supporting the operator and moving them further away from the decision-making process, it was expected that when the automation was imperfect, operator conflict detection performance would degrade because of high trust in the automation. There is concern that a failure of automation for equipped aircraft may be translated to a failure in automation for all aircraft in a mixed-equipage environment. Hence, it was hypothesized that performance in the imperfect condition as compared to the reliable condition, would further degrade because participants would be attempting to manage all aircraft. If this were the case, it was additionally expected to find increased eye movements to equipped aircraft.

Additionally, it was hypothesized that with miss prone automation: a) eye movements, visual attention, to the automated task would increase at a cost to the other tasks, and b) behavioral data would show that because the operator was paying closer attention to the raw data detection performance would improve. However, with false alarm prone automation it was expected that: a) there would be reduced fixations and dwells to the automated task, and b) behavioral performance would result in either a delayed response or no response to the automated alert.

Methods

Participants

Twelve full performance air traffic controllers (all male) aged 27 to 49 years ($M = 41$, $SD = 5.10$) volunteered and were paid \$30 per hour for their participation. Seven ATCos were from Washington, D. C., ARTCC, three were from the Washington area TRACON facility, one worked at Washington National Tower, and lastly one ATCo was from International Airport Dulles combined Tower and TRACON facilities. Participants' average work experience, including military and civilian positions, was $M = 17.5$ years, $SD = 3.63$. En route and other controllers did not differ in age ($F(1, 10) = .34$, $p > .05$) or experience ($F(1, 10) = .18$, $p > .05$).

Apparatuses: Air traffic control simulator and applied science lab eye tracker 5000

A PC-based medium-fidelity ATC simulator (Masalonis, Le, Klinge, Galster, Duley, Hancock, Hilburn, & Parasuraman, 1997) was used to simulate a generic en route airspace. The simulation consisted of a radar or primary visual display (PVD) and a data link display which were presented on two adjacent monitors. ATCos were required to monitor traffic for potential conflicts and report them by selecting the aircraft involved. A loss of separation or conflict was defined as two aircraft coming within 5nm horizontally and 1000 feet vertically of each other at all flight levels. Potential conflicts could result in an actual *conflict* or a *self-separation*. A conflict was defined as two aircraft heading towards each other at the same altitude and losing separation. A self-separation existed when one of two freely maneuvering aircraft on a conflict course made an evasive maneuver to avoid the loss of separation either by changing speed or altitude. Even though there were a total of six potential conflicts (3 self-separations and 3 actual conflicts) ATCos were instructed that they were not responsible for detecting self-separations because these aircraft had the onboard technology to safely separate themselves from the other aircraft in the sector. Therefore, the three actual conflicts were all between managed aircraft.

There were two types of aircraft in this experiment: freely maneuvering aircraft that provided ATCos with intent information and fully managed aircraft. The fully managed aircraft had yellow boxes around the datablocks to draw ATCos' attention to these aircraft, however freely maneuvering aircraft that provided ATCos with intent information did not have any boxes. The proportion of mixed-equipage was

such that 80% of the aircraft were fully managed and 20% were freely maneuvering. All freely maneuvering aircraft were above 36,000 feet and all fully managed aircraft were below 36,000 feet.

In addition to detecting conflicts, ATCos were required to accept aircraft into the sector, hand them off to the next sector as they were leaving, and perform a secondary task. The ATCos' embedded secondary task was to monitor the progress of aircraft on electronic flight progress strips. Traffic density was relatively high (about 20 aircraft controlled simultaneously) in all scenarios. An Applied Science Lab (ASL) 5000 eye-head tracking system was used to obtain eye point-of-gaze data at a sampling rate of 60 Hz. Participants were trained for one hour to ensure they were trained sufficiently with the tasks and simulation.

Design

The design was a single factor within-subjects design with four levels. The independent variable was automation support with the following four levels: 1) manual, 2) reliable, 3) miss prone automation, and 4) false alarm prone automation. Half of the participants received the manual condition first and half the participants received the manual condition last. Both conditions received reliable automation before the two unreliable automation conditions. This was done intentionally so that ATCos received sufficient experience with reliable automation in an effort to build their trust in the support tool. The ordering of the automation failure conditions was such that half of the participants performed the miss prone automation condition before the false alarm prone automation condition and the order was reversed for the other half of the participants. Thus a double crossover design was used, where the first crossover was manual versus automation conditions and the second crossover was the order of the two automation failure conditions.

Procedures

ATCos received one hour of instruction and training before performing four 30-minute scenarios. Controllers rated their mental workload, trust and self-confidence after each scenario. Following automated scenarios, ATCos rated their trust and self-confidence in the automation.

Results

Data from all twelve participants were included in the analysis. According to the hypotheses three planned

orthogonal contrasts were used to analyze if performance was affected by 1) automation versus manual control, 2) reliable versus imperfect automation support, and 3) miss prone versus false alarm prone automation. The primary question of interest was how automation imperfection affected performance. Therefore, data for all concurrent tasks and eye movements were analyzed post imperfection in an effort to avoid the "washing out" of performance differences due to averaging pre-imperfection and post-imperfection data. The exception was conflict detection performance which was analyzed at three different data points: conflict one (reliable automation for all conditions), conflict two (imperfect automation for imperfect conditions), and conflict three (reliable automation for all conditions). The F value for the omnibus ANOVA (i.e. one-way ANOVA with four levels of the independent variable) are presented first followed by the F values for the three planned orthogonal contrasts for all dependent variables as appropriate.

Primary Task Performance: Detection of Conflicts

Detection of Conflicts. Detection accuracy for the three conflicts in each condition was not submitted to the omnibus. The three planned orthogonal contrasts are reported by conflict rather than the mean conflict detection accuracy for each condition in order to avoid a masking of the effects. The data revealed no significant effects for the first, pre-imperfection conflict.

For the second conflict, ATCos detected conflicts better with reliable ($M = 100\%$, $SE = 0\%$) versus imperfect automation ($M = 37.5\%$, $SE = 14.07\%$), $F(1, 33) = 19.64$, $p < .01$. There was a nonsignificant trend for better performance with false alarm prone automation ($M = 50\%$, $SE = 15.08\%$) as compared to miss prone automation ($M = 25\%$, $SE = 13.06\%$), $F(1, 33) = 2.35$, $p = .13$. Note accuracy with miss prone automation was only half that of false alarm prone automation, it could be that the low sample size prevented this difference from reaching significance. There were no other significant results. Table 1 shows data for conflict two at all four scenarios.

For the third conflict, post imperfection, ATCos detected conflicts better with ($M = 83.33\%$, $SE = 10.88\%$) than without automation support ($M = 25\%$, $SE = 13.06\%$), $F(1, 33) = 20.13$, $p < .01$. There were no other significant effects.

| Table 1 | | | | |
|-----------|--------|----------|--------|-------------|
| Condition | Manual | Reliable | Miss | False Alarm |
| n | 12 | 12 | 12 | 12 |
| Mean | 75% | 100% | 25% | 50% |
| SE | 13.06% | 0% | 13.06% | 15.08% |
| Minimum | 0 | 0 | 0 | 0 |
| Maximum | 100 | 100 | 100 | 100 |
| Mode | 100 | 100 | 0 | Bimodal |
| Median | 100 | 100 | 0 | 50 |

Advanced Notification Time for Conflicts. Similar to conflict detection accuracy, the three planned orthogonal contrasts are reported by conflict rather than the mean advance notification time for each condition in order to avoid a masking of the effects. For the first conflict, pre-imperfection, ATCos detected the conflict earlier under automated conditions ($M = 296.07$ s, $SE = 28.75$ s) than under manual conditions ($M = 218.28$ s, $SE = 15.42$ s), $F(1, 33) = 12.71$, $p < .01$. ATCos detected the conflict earlier under imperfect ($M = 350.78$ s, $SE = 36.38$ s) versus reliable automation conditions ($M = 186.66$ s, $SE = 13.49$ s), $F(1, 33) = 46.18$, $p < .01$, reflecting an order effect. There were no other significant findings for conflict one.

For the second conflict, in 50% or more of the cells in the miss and false alarm prone conditions, no advance notification time was available due to the high number of missed events. Thus there were too few data points to carry out meaningful inferential statistics for the three contrasts.

For the third conflict, post imperfection, there were too few data points for the manual condition to perform the first contrast. There were no significant effects for the remaining contrasts.

Concurrent Tasks Performance: Accepting and Handing-Off Aircraft

Accepting Incoming Aircraft. Overall controllers accepted over 99% ($M = 99.39\%$, $SE = .51\%$) of aircraft into their sector with an average response time of 46.39 seconds ($SE = 11.13$ s). The omnibus analysis did not reveal an overall effect of automation on accuracy, $F(3, 33) = 1.32$, $p = .285$. No significant effects were found for aircraft timeliness.

Handing-Off Aircraft. Overall controllers accepted over 84% ($M = 84.82\%$, $SE = 4.81\%$) of aircraft into their sector with an average response time of 53.90 seconds ($SE = 9.92$ s). Results revealed no

significant effects for the percentage of successful hand-offs, $F(3, 33) = .686$, $p = .567$, or hand-off timeliness, $F(3, 33) = .99$, $p = .409$. However, there was an insignificant trend for better hand-off timeliness when ATCos worked with miss prone automation ($M = 86.08$ s, $SE = 3.48$ s) versus false alarm prone automation ($M = 82.17$ s, $SE = 4.23$ s), $F(1, 33) = 2.89$, $p = .098$. There were no other significant results.

Secondary Task Performance: Updating Flight Strips

Consistent with previous findings, overall performance on this embedded secondary task was extremely low ($M = 22.17\%$, $SE = 10.17\%$), perhaps due to ATCos shedding this task in an effort to maintain satisfactory performance on the main task of conflict detection. Results revealed no significant effects.

Subjective Ratings (Mental workload, trust and self-confidence)

Subjective Mental Workload Ratings. Participants indicated on a scale from 0 – 100, their perceptions of mental workload with 100 being the highest. Unexpectedly, the omnibus was not significant ($F(3, 33) = 1.43$, $p = .25$) neither were the three contrasts of manual versus automated support ($F(1,33) = 1.72$, $p > .05$), reliable versus imperfect automated support ($F(1, 33) = 2.25$, $p > .05$), and miss versus false alarm prone automation ($F(1, 33) = .31$, $p > .05$).

Subjective Trust and Self-confidence Ratings. ATCos were asked to assess their trust and reliance on the automation support tool, their impressions of how much the automation improved their performance, and lastly how self-confident they felt that they could perform without the automation. Unexpectedly, the effect of automation condition was not significant for neither the amount of trust ($F(2, 22) = .55$, $p > .05$) nor reliance ($F(2, 22) = .66$, $p > .05$) on the automation support tool. There was a significant effect of automation condition on ATCos' perceptions of how much the automation improved their performance, $F(2, 22) = 4.24$, $p = .028$. ATCos did not perceive an improvement in their performance with automated support when contrasting reliable and imperfect automation ($F(1, 22) = 1.25$, $p > .05$), however there was a marginal trend suggesting that ATCos perceived that false alarm prone automation improved their performance ($M = 69.17$, $SE = 7.73$) more than miss prone automation ($M = 53.33$, $SE = 10.01$), $F(1, 22) = 3.75$, $p = .07$. The overall omnibus testing the effect of automation on ATCo self confidence to perform

better without the automation failed to reach significance. However, ATCos were more self confident that they could perform better without the automated support in the imperfect conditions ($M = 62.5$, $SE = 7.23$) as compared to the reliable conditions ($M = 48.75$, $SE = 6.57$), $F(1, 22) = 4.72$, $p = .04$. There was an insignificant trend for increased self-confidence to perform without the automation in the miss prone scenario ($M = 69.58$, $SE = 7.80$) compared to the false alarm prone scenario ($M = 55.42$, $SE = 6.67$), $F(1, 22) = 3.76$, $p = .07$. There were no other significant results.

Eye Movements

Based on the hypothesis that miss prone automation would lead to increased visual attention (eye movements) to the automated task and false alarm prone automation would result in reduced eye movements to the automated task we investigated differences in visual attention following an automation failure (i.e. the last 16 minutes of a scenario). Eye movements up to the point when a loss of separation occurred and the automation was imperfect were excluded. It is possible that differences in visual attention strategies may have lead to the detection or non-detection of a conflict, but this was not the main interest of this study. Rather this study was concerned with attention strategies following various types of automation imperfections.

Eye Movements to Areas of Interest (Number of Fixations). The number of fixations on the different areas of interest was significantly different, $F(4, 44) = 150.02$, $p < .01$. Most fixations were made to the radar display ($M = 1209.56$, $SE = 117.48$), followed by the flights area ($M = 276.23$, $SE = 47.16$), communications ($M = 140.52$, $SE = 24.49$), resolution ($M = 55.88$, $SE = 9.74$), and flight strips ($M = 42.69$, $SE = 16.47$). Given that ATCos mostly fixated to the radar display, the remainder of the analyses focused on fixations to the radar display. The effect of the automation condition on the number of fixations to the radar display failed to reach significance, $F(3, 33) = .65$, $p > .05$, similarly none of the three contrasts were found to be significant.

Eye Movements to Areas of Interest (Duration of Fixations). The effect of the automation condition on the duration of fixations to the radar display failed to reach significance, $F(3, 33) = 1.32$, $p > .05$, similarly none of the three contrasts were found to be significant.

Eye Movements to Areas of Interest (Number of Dwells). The effect of the automation condition on

the number of dwells on the radar display was marginally significant, $F(3, 33) = 2.50$, $p = .076$. There was an increased number of dwells on the radar display with imperfect automation ($M = 87.92$, $SE = 6.76$) as compared to reliable automation ($M = 66.92$, $SE = 7.9$), $F(1, 33) = 6.60$, $p = .015$.

Eye Movements to Areas of Interest (Total Dwell Time). The effect of the automation condition on the total dwell time to the radar display failed to reach significance, $F(3, 33) = 1.347$, $p > .05$, similarly none of the three contrasts were found to be significant.

Eye Movements to Aircraft Projected to be in Conflict (Optimal Distribution of Attention by Type of Conflict). We computed a measure of optimal distribution of attention (ODA). Our measure of ODA was based on the assumptions that better attention distribution was obtained by better situation awareness as evidenced by 1) an even distribution of fixations to the two aircraft in conflict and 2) a high number of fixations to the two aircraft in conflict ($ODA = \text{difference of fixations} / \text{sum of fixations}$). We computed the sum and difference between the numbers of fixations to the two aircraft in conflict. The rationale is that if controllers evenly distributed attention (roughly equal number of fixations) to the two aircraft in conflict, then they may have better situation awareness regarding the conflict pair. Conversely, if the difference was large their situation awareness regarding the two aircraft in conflict may be lower. Further, by dividing the difference by the sum we were able to get a measure relative to the total number of fixations for each conflict pair. Better ODA was evidenced by the ratio (sum of fixations / difference of fixations) approaching zero.

For conflict one, pre-imperfection, there was an insignificant trend for automation condition on ODA, $F(3, 33) = 2.2$, $p = .11$. ODA was better with reliable automation ($M = .83$, $SE = .11$) as compared to the imperfect automation conditions ($M = 1.0$, $SE = 0$), $F(1, 33) = 5.87$, $p = .02$. There were no other significant findings.

For conflict two, the omnibus was insignificant $F(3, 33) = 1.78$, $p = .17$. However, ODA was better for miss ($M = .45$, $SE = .09$) versus false alarm prone automation ($M = .69$, $SE = .09$), $F(1, 33) = 5.26$, $p = .028$. The data revealed no other significant results.

Following the imperfection, there was a significant effect of automation for conflict three, $F(3, 33) = 4.60$, $p = .009$. ODA was better under manual control ($M = .38$, $SE = .09$) compared to automated support ($M = .66$, $SE = .08$), $F(1, 33) = 10.32$, $p = .003$. There was a

marginal trend for better ODA for miss ($M = .55$, $SE = .11$) as compared to false alarm prone automation ($M = .79$, $SE = .21$), $F(1, 33) = 3.26$, $p = .08$. There were no other significant findings.

Discussion

The purpose of this experiment was to investigate ATCo performance, mental workload, and attention allocation with reliable and imperfect automation support in a mixed-equipage environment. Similar to previous findings (Metzger & Parasuraman, 2005), the hypothesis that ATCos would detect more conflicts and detect conflicts sooner when they were supported by reliable automation as opposed to performing the tasks manually was upheld with the exception that we did not see performance differences for the concurrent tasks.

When the automation was imperfect, operator conflict detection performance degraded catastrophically, thereby supporting previous work with automation imperfection in an ATM environment (Metzger & Parasuraman, 2005). This suggests that automation supports the operator, yet moves them further away from the decision-making process so that when the automation is imperfect operators are caught over-trusting the automation resulting in poor performance (Metzger & Parasuraman, 2001). Even though ATCos made an increased number of dwells to the radar display when they were supported with imperfect automation as compared to reliable automation, conflict detection performance still degraded. The increased number of dwells to the radar display with imperfect automation may explain why the typical eye movement related affects of complacency were not found, i.e. reduced visual attention with versus without automation support. A simple inspection of the data showed a decreased number of fixations with reliable automation support ($M = 66.92$, $SE = 7.9$) as compared to manual control ($M = 78.38$; $SE = 8.86$).

Detection of the reliably cued conflict suffered when another conflict pair was incorrectly announced (false alarm event) within the same time period. Similar to previous work (Meyer, 2001; Dixon & Wickens, 2006), it seems that the cry wolf phenomenon was upheld such that 50% of the ATCos failed to respond to the reliably cued conflict when a false alarm for another conflict pair was present. An alternative explanation is that ATCos may have experienced a significant amount of workload and consequently could not respond to the reliably cued event because they were busy cross verifying the false alarm event. However,

neither an increase in eye movements nor ratings of subjective workload support this explanation.

While conflict detection performance following a miss and false alarm was directly compared by evaluating the detection of the same conflicts, our hypothesis of improved conflict detection with miss prone and delayed or no response to false alarm prone automation was not upheld in the third conflict. This could be because when ATCos failed to detect a conflict regardless of the type of automation imperfection, they were provided with feedback in the form of red warning circles appearing around the aircraft involved in the loss of separation aircraft.

It is very encouraging that following feedback, ATCos' conflict detection performance for both miss and false alarm prone automation improved dramatically. Poor performance in a time critical environment with imperfect automation raises safety concerns, even more so, it is striking that the cry wolf phenomena was upheld with false alarm prone automation in the face of a nearly concurrent reliably cued event for the primary task of conflict detection.

References

- Corker, K., Fleming, K., & Lane, J. (1999). Measuring controller reactions to free flight in a complex transition sector. *Journal of ATC*, 4, 9-16.
- Dixon, S. R., & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle flight control: Evaluating a reliance-compliance model of automation dependence in high workload. *Human Factors*, 48, 474-486.
- Masalonis, A. J., Le, M. A., Klinge, J. C., Galster, S. M., Duley, J. A., Hancock, P. A., Hilburn, B. G., & Parasuraman, R. (1997). Air traffic control workstation mock-up for Free Flight experimentation: Lab development and capabilities. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting (pp. 1379). Santa Monica, CA: HFES
- Metzger, U. & Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management: An empirical study of active vs. passive monitoring. *Human Factors*, 43, 519-528.
- Metzger, U., & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors*, 47(1), 35-49.
- Metzger, U., Rovira, E., & Parasuraman, R. (2003). Controller performance, workload and attention allocation in distributed air-ground traffic management: Effects of mixed equipage and decision support. *Proceedings of the 12th International*

Symposium on Aviation Psychology. Columbus, OH: Ohio State University.

Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors*, 43, 563-572.

Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors*, 46, 196-204.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced "complacency". *The International Journal of Aviation Psychology*, 3, 1-23.